



Alanine Scanning Mutagenesis Of A High-Affinity Nitrate Transporter Highlights The Requirement For Glycine And Asparagine Residues In The Two Nitrate Signature Motifs

By: Unkles, S; Karabika, E; Symington, V; **Cecile, J**; Rouch, D; Akhtar, N; Cromer, B, and Kinghorn, J

Abstract

The high-affinity Major Facilitator Superfamily (MFS) nitrate transporter protein NrtA has proved thus far to be intractable to crystallographic procedures. Instead, structural models were developed and are presented here based on the crystal structures of the MFS members phosphate/glycerol-3-phosphate antiporter, GlpT and fucose/proton symporter, FucP. Alanine-scanning substitution of residues within and around the two nitrate signatures (NS1 and NS2) of NrtA have been analysed and show that residues within NS2 (TM11) are more tolerant of substitution by alanine than those in NS1 (TM5). In models of NrtA, asparagine residues N168 in NS1 and N459 in NS2 are positioned approximately midway between the outside and the inside faces of the protein at the central pivot point of the protein. N168 and N459 are in close proximity to substrate-binding arginine residues R368 and R87, respectively, which lie offset from the pivot point towards the cytoplasmic face. The N168/R368 and N459/R87 residue pairs are relatively widely separated on opposite sides of the likely substrate translocation pore. Results demonstrate the critical structural contribution of several glycine residues in each NS. These glycines are at sites of close helix packing and are likely to be important for protein folding and flexibility of TMs 5 and 11 as well as positioning of N168 and N459 correctly to allow proximity of the N168/R368 and N459/R87 residue pairs. Given the relative locations of N168/R368 and N459/R87 residue pairs, the validity of the models and possible role of the NSs together with the substrate-binding arginine residues are discussed.

Introduction

Nitrate is a relatively simple small inorganic molecule with an ionic radius of only 1.96 Å, but has a dual character of crucial considerable biological positivity or negativity. As Dr. Jekyll, nitrate acts a major nitrogen source in natural environments for numerous bacteria, fungi, algae, plants including crop plants and it is the nutrient that most frequently limits their growth (see reviews 1,2, and references therein). In addition, vast quantities of nitrogen-based fertilizers, rapidly converted to nitrate by soil oxidising bacteria (3) are widely used in agriculture to sustain maximum crop yields and thereby satisfying the needs of our increasing human population and its food demands. Nitrate in this beneficial mode, is assimilated to ammonium and catalytically transformed to amino acids for protein synthesis and growth of these organisms. Furthermore, vegetable dietary nitrate in the human oral cavity and gut may be utilized by nitrate reducing symbiotic bacteria to provide protection against pathogenic bacterial species (4,5, and references therein). In Mr. Hyde mode, nitrate is a serious water pollutant readily leaching from soils, contributing to eutrophication of natural water systems with resultant loss of biodiversity and simultaneously lowering nitrogen availability to environment land as well as crop plants (6). Second a derivative of nitrate, nitrous oxide, emitted from both anthropogenic and ecology activities, is a major greenhouse gas (7). Finally, there are human and animal health concerns as carcinogenic N-nitrosamines also derived from nitrate may be generated by bacteria in the human gut (4, and references therein).

The transport of nitrate across membrane barriers by several classes of proteins, including the ABC transporters (8), the low-affinity transporters (9) and the high-affinity transporters (10), has been studied by a number of research groups. The high-affinity nitrate transporters are particularly important

being present in all kingdoms except animal. Following the description of the first gene encoding a high-affinity nitrate transporter from *Aspergillus nidulans* (11), it became clear that such transporters belong to the nitrate/nitrite porter family, NNP (12,13) a subfamily forming a distinct cluster of the largest secondary transporter family, the major facilitator superfamily (MFS: TC 2.A.1). This superfamily comprises a wide range of functionally diverse transporters (14), most of which possess 12 transmembrane α -helices (TMs). Whilst the determination of the structure of a high-affinity nitrate transporter has proven recalcitrant to crystallography, the structures of several prokaryotic MFS transporters, *vis-à-vis* the lactose transporter, LacY (15), the phosphate/glycerol-3-phosphate antiporter, GlpT (16), the multidrug transporter, EmrD (17), the fucose transporter, FucP (18) and the peptide transporter, PepT_{So} (19) have been solved. However, although the location of residues and motifs are much more accurately mapped from the 3D crystal structure, than from 2D secondary structure models, the transport mechanisms involved may still be obscure without detailed biochemical information.

Previous studies of *Aspergillus nidulans* NrtA (and *E. coli* NarK) have indicated that two perfectly conserved arginine residues, R87 and R368 in TM2 and TM8, respectively, are essential and most likely represent the substrate binding site for nitrate in the NNP family proteins (20,21). In addition, second site suppressor results indicated that at least one of these, R87, was located in close proximity to an asparagine residue, N459, in TM11. N459 resides in a region of the transporter that is recognised as a conserved motif present in all NNP family transporters, and is not observed in any other class of MFS proteins (13). This residue stretch has been referred to as the nitrate signature (NS) of which there are two copies within NrtA, located in TM5 and TM11 (Fig. 1).

Towards understanding the importance of NS1 and NS2 and their role in the structure and function of this anion transporter, we have carried out systematic alanine scanning *in-vitro* mutagenesis on these motifs. Specifically, we

have altered NS1 and NS2 amino acids, as well as several residues N- or C-terminal to NS1 and NS2, to alanine and biochemically characterised the resulting mutant strains. In addition, we have developed models for the nitrate transporter based on the crystal structures of GlpT (inward-facing) and FucP (outward-facing) (16,18) to complement the analysis of the mutant information, as well as allowing a re-evaluation of previous mutant data (20) and assessment of possible mechanisms for the transport process by NNP proteins. Our analysis indicates that the nitrate signatures fulfil several functions in nitrate transporters – conserved glycine residues are required for positioning of helices as well as for close helix packing and flexibility, while N168 in NS1 and N459 in NS2 may have a more direct role in nitrate transport.

EXPERIMENTAL PROCEDURES

Fungal Strains - Aspergillus nidulans strains used in this study were (a) the standard wild-type (with regard to nitrogen metabolism) the biotin auxotroph, strain *biA1*, (b) the double deletion mutant *nrtA747 nrtB110* (disrupted in both *nrtA* and *nrtB*), strain T110 and (c) triple mutant strain harbouring *nrtA747 nrtB110* mutations as well as the arginine auxotrophic marker *argB2*, strain JK1060.

Aspergillus nidulans Media - Standard *Aspergillus* growth media and handling techniques were as described before (22). Shake flask cultures for nitrate uptake assays were grown in liquid minimal medium (23, and as modified by 24).

Escherichia coli Strains, Plasmids and Media - Standard procedures were used for propagation of plasmids, as well as for subcloning and maintenance of plasmids within *E. coli* strain DH5 α .

Genetic Transformation Procedure - The *A. nidulans* transformation procedure to obtain single copy integration at the *argB* locus, was essentially that carried out as described before (reviewed in 25, and references therein). For each mutation, around ten arginine auxotrophic transformant strains were analysed by Southern blotting to identify a single *nrtA* gene copy integration at *argB*.

Molecular Analyses - Generation of amino acid replacement constructs by PCR overlap extension was as described previously (26) with all constructs being verified by DNA sequencing. Constructs all included a sequence resulting in C-terminal fusion of a V5 tag (27) to the NrtA protein. *Bam*HI digestions of genomic DNA were analysed by Southern blotting following transformation of constructs as described previously (20). For PCR amplification to verify the sequence of single copy alanine replacement mutants, DNA was isolated from 50 mg pressed wet weight of mycelium suspended in 500 μ L of breaking buffer (2% Triton X-100, 1% SDS, 100 mM NaCl, 10 mM Tris-HCl, pH 8.0, and 1 mM EDTA, pH 8.0), to which was added 500 μ L of chlorophorm:isoamyl alcohol (24:1) and 300 μ L of 0.5 mm glass beads (BioSpec). Mycelium was homogenised using a FastPrep 24 homogeniser at speed 4.5 for 20 sec followed by centrifugation at 16,000 g for 10 min at room temperature. The upper phase was transferred to a new 1.5 mL microcentrifuge tube, two volumes of ice cold 96% (v/v) ethanol added and mixed. DNA was pelleted by centrifugation at 16,000 g for 10 min, washed with 70% (v/v) ethanol, dried and resuspended in 50 μ L TE buffer containing 10 μ g/mL RNase A. 0.5 μ L of the DNA was used in 50 μ L PCR reactions using Phusion High Fidelity DNA polymerase (NEB) and primers MUTF and MUTR2.

Bioinformatics – Sequence logos were generated from an alignment of 500 prokaryotic and 173 eukaryotic proteins identified in a BlastP search using the *Escherichia coli* NarU sequence as query. Multiple alignments were generated using MAFFT (26) at the European Bioinformatics Institute (www.ebi.ac.uk) and entered into Jalview (29) from which required sequences were extracted and entered into Weblogo (weblogo.berkeley.edu). Color information was discarded from the Weblogo output and logos were presented in greyscale.

Net Nitrate Uptake Assays – These were performed on cultures grown in minimal medium with 5 mM urea as nitrogen source for a total of 6.5 h at 37 °C with induction of the

transporter by addition of 10 mM sodium nitrate 100 min prior to assay, as described previously (30). Assays measuring depletion of nitrate from the medium after incubation for 20 min at 37 °C were carried out on at least three independently grown cultures for each strain and expressed as nmol nitrate removed from the medium min⁻¹ mg⁻¹ dried mycelium.

Protein Expression Analysis - For immunological detection, cultures were grown as for the net nitrate uptake assays and crude membrane preparations made from 50 mg pressed wet weight of mycelium. Mycelium was suspended in 500 µL ice cold extraction buffer consisting of 10 mM sodium orthophosphate, 200 mM sodium chloride, 10% (v/v) glycerol, pH 7.0 to which was routinely added fresh 0.1 mM PMSF and 1 mM benzamidine. As mentioned in the Discussion section, some preparations were tested with additional complete protease inhibitor cocktail (Roche) in the extraction buffer. 300 µL of 0.5 mm glass beads was added and cells disrupted by homogenisation at speed 4.5 for 20 sec in a FastPrep 24 homogeniser. After 2 min on ice, samples were centrifuged at 16,000 g for 1 min at room temperature. The supernatant fraction was centrifuged at 4 °C for 45 min at 25,000 g and the resulting crude membrane pellet resuspended in 50 µL cold extraction buffer. Samples of 1 µL, equivalent to around 50 µg protein, were run on a 4-12 % NuPAGE gel (Invitrogen), blotted on to Invitrolon membrane (Invitrogen) as described by the manufacturer and NrtA protein detected using anti-V5 HRP conjugated antibody (Invitrogen) and an ECL Plus Western Blot Detection System (GE Healthcare). At least three independently grown cultures were analysed for each strain and representative results are shown.

Generation of 3D NrtA models – Due to low sequence homology between MFS proteins, molecular structural models for NrtA were built using threading, or fold recognition, servers Phyre (31) and iTasser (32). Models were returned based on MFS templates, including the phosphate:glycerol-3-phosphate antiporter, GlpT, the lactose permease, LacY,

the fucose transporter, FucP, and the oligopeptide transporter, PepT_{so}. The similarity in the models based on different templates and from the two different servers gave confidence that the results were credible models. The models based on GlpT (PDB ID: 1PW4) and FucP (PDB ID: 3O7Q) were selected as representatives of the open-inward and open-outward states, respectively. The large intracellular loop between TMs 6 and 7 from T240 to F291 was removed from these models. Pymol (www.pymol.org) software was used to analyse the models and produce structural cartoon figures.

RESULTS

Amino acid sequence comparative analysis of NS1 and NS2 – Fig. 2 provides a graphical display of sequence conservation generated from multiple amino acid sequence alignments obtained from putative nitrate transporters of prokaryotes and eukaryotes where the height of a residue reflects its frequency at that position. Within these displays can be seen the sequences originally identified as the nitrate signature motifs of NrtA (12,13). However, with the increasing availability of genome sequences and the consequent rise in predicted nitrate transporter proteins, such alignments permit a reassessment and redefinition of the most highly conserved residues in and around NS1 and NS2. Clearly from Fig. 2, the prokaryotic signatures show a higher level of sequence divergence (i.e. greater ‘choice’ of residues) than eukaryotes at many of the positions within the alignment, possibly as a consequence of the far larger number of prokaryotic sequences available, with conservation in NS1 being stronger than NS2. Nevertheless, certain residues appear almost invariant such as the glycine at position 1 of the alignment, glycine at position 10 of NS2 and the GNXG at positions 11 to 13 of NS1. These latter residues and the position 1 glycine are also highly conserved in the eukaryotic signatures as well as others such as asparagine at position 4 of NS1 and glycine at position 5 of NS2, glycine at position 8 in both signatures and glycine residues at positions 15 and 16 of NS1.

3D models of NrtA - The actual sequences of the NrtA signatures relative to the alignment positions are shown in Fig. 2. Thus, glycine 1 of the alignment is equivalent to G157 of NS1 and G448 of NS2, and these highly conserved residues from the eukaryotic sequences in Fig. 2 can be observed in the NS1 and NS2 sequences of NrtA. The 3D models of NrtA in Fig. 3 are based on the crystal structures of the open-inward GlpT (Fig. 3A, C and E) and open-outward FucP (Fig. 3B, D and F), conforming to the alternating access mechanism. Residues G157 and G448 locate near the N-terminus of their respective TMs 5 and 11. Directly above these in the helices are S161 and G165 in TM5 and G452 and G456 in TM11. In the open-outward, FucP-based model (Fig. 3B and D), these residues are tightly packed against other TMs in the 'closed' portion of the molecule. Facing the lumen within the protein in which the nitrate-binding residues R87 and R368 are located (Fig. 3A-D), the first turn of TM5 contains N160 that is perfectly conserved in the eukaryotic NS1 sequences. Two helix-turns above this lies the N168 residue of the nitrate signature core GNXG motif (Fig. 3A and B). In TM11, the highly conserved G452 is in an approximately equivalent position to N160, two helix turns below the core motif asparagine, N459 (Fig. 3C and D). Other conserved glycine residues are placed around the helix close to N168 in TM5 (G167, G170, G171, G172) and N459 in TM11 (G455, G458, G461, G462) (Fig. 3A-D), while relatively bulky side chains such as L166 in TM5, and F457 and L460 in TM11 are on the opposite face of the helix from N168 and N459, respectively. The models in both inward- and outward-facing conformations show the position of N168 and N459 approximately central in the protein with R87 and R368 roughly two-thirds towards the cytoplasmic side (Fig. 3A-D). Significantly, within the central lumen, both models show close proximity of N168 to R368 and N459 to R87 (Fig 3E and F), the latter consistent with the functional replacement reported earlier (20).

Alanine scanning of Nitrate Signature I - Residues G157 to T174 of NS1 in TM5 were individually changed to alanine,

including residues of the first nitrate signature *per se*, as well as several residues flanking the signature. Natural alanine residues at positions 159, 163, 164 and 169 were left unaltered. Alanine replacement mutant strains were examined for (i) growth ability on nitrate as the sole nitrogen source, (ii) net nitrate transport where growth was observed on nitrate and (iii) protein expression levels by western blot analyses.

(i) Growth phenotypes - Alanine-scanning of all four highly conserved residues within NS1, G165A, G167A, N168A and G170A resulted in complete loss of the ability to grow on nitrate at 37 °C, indicating loss-of-function (Table 1). Similarly alanine replacement of residues conserved in eukaryotes only, permitted slight growth (G171A) whilst G172A failed to grow. In contrast, mutants with alanine replacements of residues flanking conserved residues and mainly poorly conserved, residues T158A, N160A, S161A, L162A (N-terminal to the NS1 core) as well as I173A and T174A (C-terminal to the NS1 core), retained wild type or near wild type growth. Finally, L166A (located in the NS1 core) showed intermediate growth and G157A (at the N-terminus of TM5) led to loss-of-function as judged by growth response. Growth responses were also tested on nitrate as sole nitrogen source at 25 °C to assess the possibility of mutant temperature sensitivity. For most mutants, growth at 25 °C paralleled that at 37 °C, with the exception of mutant G171A, which grew well relative to the wild type at 25 °C despite showing very poor growth at 37 °C (data not shown).

(ii) Net nitrate uptake activity - Assays were carried out for mutants that showed growth on nitrate (Table 1). As expected from the growth phenotypes, net nitrate uptake activity of the mutants T158A, S161A, L162A, I173A and I174A, located before and after the core signature, was similar to the wild type level. The exception to this was the strain possessing alanine substitution of residue, N160, a very highly conserved residue in the eukaryotic nitrate transporters. This mutant showed a notable decrease in activity compared to the wild type. Change of core signature residue G171 to alanine, resulted in

substantially reduced net nitrate uptake activity correlating with the poor growth of this mutant. Mutant L166A, concomitant with slightly better growth on nitrate than G171A, displayed an intermediate rate of nitrate uptake. Mutant G165A was chosen for assay as a representative of those mutants unable to grow on nitrate to check the level of activity was indeed as low as the negative control strain, T110.

(iii) Protein expression levels - Western blot analyses showed that the wild type strain, T454 (Fig. 4), has a single band of 45 kD whilst no protein was detected in the negative control strain, T110 (Fig. 4B). Virtually no NrtA protein was observed in crude membranes prepared from mutants in the highly conserved residues G157A, N168A, and G172A (Fig. 4A). Strains G165A, G167A and G170A showed much reduced levels of NrtA and possessed a second band of around 41 kD. While the total amount of NrtA protein detected in the crude membranes from mutants L166A and G171A was higher than for G165A, G167A and G170A, both bands of 45 kD and 41 kD were observed. The smaller band of 41 kD is suggestive of proteolysis but this is unlikely to be due to digestion during protein isolation since experiments in which a protease inhibitor cocktail was added to the extraction buffer made no difference to the appearance of the preparation in western blots (data not shown). Nor is the smaller band likely to be the result of inherent protein instability causing aberrant migration in the gel as trial experiments on the NrtA protein solubilised from crude membranes of F457A to the action of trypsin showed no increased susceptibility over the wild type protein (data not shown). Therefore, proteolysis of the mutant proteins most probably occurs during trafficking to the membrane or following insertion into the membrane. Mutants N-terminal to the core (T158A, N160A, S161A, L162A) and mutants C-terminal (I173A and T174A) possessed a single wild type band of around 45 kD.

Alanine scanning of Nitrate Signature
2 - Residues G448 to L465 in TM11 covering the NS2 core region as well as flanking regions

was altered to alanine and mutant strains were examined as for NS1 mutants.

(i) Growth phenotypes - In contrast to many of the mutants in the NS1 region, most alanine substitution mutants in the NS2 area retained the ability to grow on nitrate at 37 °C as a sole nitrogen source (Table 1). Only N459A completely lacked growth. Reduced growth on nitrate relative to the wild type was observed for mutants of highly conserved residues, G448A and G452A, and of those within the NS2 core region, G455A, G456A, F457A, G458A, L460A, G461A and G462A. Other alanine substitution mutants, I449A, V450A and S451A, N-terminal to the core residues, and I463A, I464A and F465A, C-terminal to the core residues, showed growth comparable to the wild type. At 25 °C, most mutants displayed a similar pattern of growth on nitrate as at 37 °C, except for mutants G448A and G455A which appeared to grow as well as the wild type strain at 25 °C (data not shown).

(ii) Net nitrate uptake activity - Concomitant with the growth phenotypes, net nitrate uptake was observed for most of the mutants in the NS2 region, with those displaying wild type growth on nitrate having uptake rates around 8 nmol nitrate/min/mg, similar to the wild type (Table 1). Intermediate values of between 1.2 and 4.7 nmol nitrate/min/mg were obtained for those mutants that grew less well on nitrate.

(iii) Protein expression levels - Mutants flanking the core region, I449A, V450A, S451A, M453A, V454A (N-terminal) and I463A, I464A and F465A (C-terminal) showed a single band of 45 kD similar to the wild type (Fig. 4B). Mutants in the core region showed, in some cases, in total almost wild type protein levels (G455A, G456A, F457A, G458A, G461A, and G462A) or reduced expression (L460A) and all possessed a second band of around 41 kDa. Similarly, crude membranes of mutants of highly conserved residues G448 and G452 altered to alanine showed good expression (G448A) or poor expression (G452A) and two bands of 45 and 41 kD. Almost no NrtA expression was observed in strain N459A.

DISCUSSION

The clear deleterious effects of alanine scanning mutagenesis on residues within the nitrate signatures underline the importance of these conserved motifs as structural features necessary to facilitate the structure and/or function of the NrtA protein. It is interesting that most of the residues showing severely reduced protein expression occur in NS1 suggesting that, although apparent symmetry partners in the structural models, TM5 is less able to accommodate change than TM11. This may simply be due to stability afforded by the N-terminal half of the protein during folding of the C-terminal portion. However, the asymmetry in the effect of mutations is also consistent with the functional asymmetry noted previously for LacY (33) and NrtA (20) as well as the recent suggestion from interpretation of the crystal structure of PepT_{so} that the N-terminal six helices may be less dynamic than the corresponding C-terminal helices (19). The more frequent retention of activity and protein expression of alanine substitutions in NS2 compared to NS1 may therefore reflect less rigid structural constraints in NS2.

A notable feature of the nitrate signatures is the large number of glycine residues. Glycine residues offer greater conformational flexibility (bend and swivel) than other side chains and, whilst known as helix-breakers in soluble proteins, occur commonly in transmembrane helices, often involved in close helix-helix contacts (34). Both nitrate signatures contain serial sequences of glycine residues, two or three together, spacing that suggests flexibility and/or rotation of the strand may be required for optimal function. Indeed, molecular dynamics simulations have highlighted the highly flexible nature of TMs 5 and 11 in both LacY and GlpT (35,36). In NrtA, the high incidence of glycine residues in the nitrate signatures suggests a particular requirement for flexibility. The larger side chain of alanine compared to glycine substantially reduces the possible range of bending and/or rotation of a polypeptide strand (37). Thus the observation that substitution of glycine by alanine in the equivalent TMs of NrtA leads to proteins of low or no activity, along with poor folding of

the whole protein (as evidenced by the 41 kDa band observed in western blots of several mutants) is consistent with the need for structural flexibility exhibited in the LacY and GlpT simulations.

In addition to flexibility, a key property of glycine residues located within TMs is that their small side chain allows close helix-helix interaction, exemplified by the GXXXG motif of glycophorin A and other proteins (38,39). In OxIT and GlpT, GXXXG/S motifs are present in TMs 5 and 11 towards the cytoplasmic end of the helices. It was postulated that in OxIT these motifs are required during the alternating access process to aid packing of the helices that close the cytoplasmic side of the protein when in the conformation open to the periplasm (40). Recent molecular dynamic simulations of GlpT showing closure of the cytoplasmic lumen by movement of TMs 5 and 11 support this proposition (41). In NrtA, two contiguous and overlapping GXXXG/SXXXG motifs are present at the cytoplasmic ends of both TM5 and TM11 at residues 157-165 in TM5 and residues 448-456 in TM11. These are reminiscent of glycine zipper motifs (42) and, as expected for such motifs, are well conserved in nitrate transporters (Fig. 2A), with perhaps the exception of TM11 of the prokaryotic proteins. In nitrate transporters, it is likely therefore that the contiguous cytoplasmic GXXXG/SXXXG motifs are required for tight helix packing to close the cytoplasmic end of the protein in the outward-facing conformation. In the open-outward model (Fig. 3B and D) glycine residues of the GXXXG/SXXXG motifs all reside on the same face of the helix as the asparagine residues N168 and N459 in their respective TMs in an orientation suggesting directional close packing of TM5 towards TM8 (the location of substrate-binding residue R368) and TM11 towards TM2 (the location of substrate-binding residue R87). The preponderance of glycine should allow close and dynamic association of the nitrate-binding residues R87 and R368 with N459 and N168, respectively. Indeed, in the helix turn immediately C-terminal (toward the cytoplasmic face) to R87 in TM2 and R368 in

TM8 are one or two, respectively, very highly conserved glycine residues (G91 in TM2, G371 and G372 in TM8) whose location could allow very close proximity of N459 and N168, respectively.

In GlpT, another GXXXG motif near the centre of TM5 of OxIT was thought to be less important for inter-helix association (40). In NrtA, GXXG motifs encompass the conserved asparagine residues N168 and N459, roughly in the centre of TM5 (residues 167-170) and TM11 (residues 458-461) (Fig. 3. *A-D*). In both the inward-facing and outward-facing models of NrtA, the glycine residues in these motifs are positioned such that those in TM5 about TM1 and those in TM11 are closest to TM7, just at the points of crossover of TM5 with TM1 and TM11 with TM7. These GXXG motifs are well conserved and often immediately followed by up to three residues of small side chain volume, principally glycine residues in eukaryotic nitrate transporters (Fig. 2). We propose that these central GXXG motifs at positions 167-170 in TM5 and 458-461 in TM11, and their associated glycine residues G171 and G172 (TM5) and G462 (TM11), are critical to inter-helix crossover thus permitting juxtaposition of helices carrying the residues for substrate transport.

Prominent in both NS1 and NS2 are the conserved asparagine residues that have previously been shown to be irreplaceable (N168) or to allow only very conservative change to glutamine (N459) (20). Both preceded and followed by glycine residues are L166 (on the opposite face of the helix from N168) and F457 (on the opposite face from N459). With analogy to the proposed potassium channel gating hinge mechanism (43), these relatively large residues could provide a hinge, compression of which permitted by the adjacent glycine residues, would propel the respective asparagines residues into the substrate translocation pathway. Change of L166 or F457 to alanine might not substantially alter this function and permit the observed level of activity. Proximity of N459 with the perfectly conserved arginine, R87 in TM2, previously inferred by complementary substitutions (20), can be seen in the models of NrtA as can the equivalent

positioning of N168 close to R368 (Fig. 3). Given the evidence indicating that R87 and R368 are substrate-binding residues (20), the proximity of the asparagine residues in the nitrate signatures to the respective arginine side chains alludes to a potentially important function in nitrate transport for N168 and N459. The complete loss of detectable protein by western blot for the two mutants N168A and N459A, demonstrates that these residues are critical for a stable folded protein structure. From the position of the residues in the models shown in Fig. 3, we can infer that the structural role of these two asparagine residues is in mediating interactions across the interface between the N- and C-terminal halves of the protein, right at the apparent location of the hinge between the two alternating access conformations. However, because of this important structural role, we cannot conclude anything from our data about a specific functional role for these two asparagines in the signature motif. Nevertheless, with their propensity for hydrogen bond formation and their location, it seems likely that these residues have in addition a function in nitrate coordination.

One common theme in MFS anion transporters is the requirement for two positively-charged residues in the coordination of negatively-charged substrates. In the best-studied examples, namely GlpT and OxIT, these positively-charged residues are located at or close to the proposed pivot point (according to the model of alternating access) of the transporters roughly at the centre of the protein. Thus, the GlpT antiporter possesses two arginine residues, R46 and R269 located in TM1 and TM7, respectively, which have been proposed to interact directly with phosphate or the phosphate moiety of glycerol-3-phosphate (16). Molecular dynamics simulations indicate that this binding by the arginine residues is not simultaneous but occurs initially to R46 and probably sequentially to R269 during the substrate translocation process (41). Combined biochemical evidence and structural modelling (based on the GlpT structure) predict that in the OxIT protein, R272 and K355 in TMs 8 and 11, respectively, form the binding site for the substrate for oxalate (44,45). Therefore, in

each of these proteins the anion-binding residues are positioned within 10 Å of each other such that the side chains can interact with substrate, either simultaneously or in sequence, effectively forming the substrate-selective binding site. NrtA likewise has two positively-charged residues, R87 and R368, implicated in nitrate binding. However, in contrast to GlpT and OxlT, relative spacing of residues based on the inward-open (GlpT) and outward-open (FucP) models clearly shows the comparatively wide separation of around 15 Å between R87 from R368. This distance is far greater than would support the simultaneous coordination of the small anion, nitrate, with an ionic radius of just 1.96 Å (46). Similarly, the residue side chains are modelled such that sequential binding would also appear improbable with this separation. In addition, from the models of NrtA, the positions of R87 and R368 in the protein are not central, but offset towards the cytoplasmic face. Undoubtedly, the distances obtained from models based on templates with low amino acid sequence similarity, such as that of NrtA for GlpT or FucP, will be subject to inaccuracies in assigning precise positions for individual residues. Nevertheless, experimental data obtained with numerous MFS transporters of diverse substrates such as the phosphate transporter, Pho84 (47), the multidrug transporter, MdfA (48), the human organic anion transporter 1, hOAT1 (49), the monocarboxylate transporter, Jen1p (50) and the sialic acid transporter, Sialin (51) have shown broad agreement with the respective models based on GlpT and supported the overall structural conservation within the MFS. In order to maintain this overall structure but retain a single substrate binding site, the close packing afforded by the presence of glycine residues conserved in nitrate transporters may allow the substrate-binding arginine side chains to be substantially closer (around 7 Å) than the basic modelling suggests. NrtA and other nitrate transporters would therefore be markedly more compact proteins than hitherto studied MFS transporters. Alternatively there may be an active intermediate structure where conformational change brings R87 and R368 transiently close enough together to bind a

nitrate ion simultaneously or to pass nitrate from one to the other in sequence.

REFERENCES

1. Fernandez, E., and Galvan, A. (2008) Nitrate assimilation in *Chlamydomonas*. *Eukaryot. Cell* **7**, 555-559
2. Dechorgnat, J., Nguyen, C.T., Armengaud, P., Jossier, M., Diatloff, E., Filleur, S., and Daniel-Vedele, F. (2011) From the soil to the seeds: the long journey of nitrate in plants. *J. Exp. Bot.* **62**, 1349-1359
3. Kowalchuk, G.A., and Stephen, J.R. (2001) Ammonia-oxidizing bacteria: A model for molecular microbial ecology. *Ann. Rev. Microbiol.* **55**, 485-529
4. Lundberg, J.O., Weitzberg, E., Cole, J.A., and Benjamin, N. (2004) Nitrate, bacteria and human health. *Nature Rev. Microbiol.* **2**, 593-602
5. Hord, N.G., Tang, Y., and Bryan, N.S. (2009) Food sources of nitrates and nitrites: the physiologic context for potential health benefits. *Am. J. Clin. Nutr.* **90**, 1-10
6. Power, J.F., Wiese, R., and Flowerday, D. (2001) Managing farming systems for nitrate control: a research review from management systems evaluation areas. *J. Environ. Qual.* **30**, 1866-1880
7. Liu, L., and Greaver, T.L. (2009) A review of nitrogen enrichment effects on three biogenic GHGs: the CO₂ sink may be largely offset by stimulated N₂O and CH₄ emission. *Ecol. Lett.* **12**, 1103-1117
8. Koropatkin, N.M., Pakrasi, H.B., and Smith, T.J. (2006) Atomic structure of a nitrate-binding protein crucial for photosynthetic productivity. *PNAS* **103**, 9820-9825
9. Tsay, Y.F., Chiu, C.C., Tsai, C.B., Ho, C.H., and Hsu, P.K. (2007) Nitrate transporters and peptide transporters. *FEBS Lett.* **581**, 2290-2300
10. Glass, A.D., Britto, D.T., Kaiser, B.N., Kinghorn, J.R., Kronzucker, H.J., Kumar, A., Okamoto, M., Rawat, S., Siddiqi, M.Y., Unkles, S.E., and Vidmar, J.J. (2002) The regulation of nitrate and ammonium transport systems in plants. *J. Exp. Bot.* **53**, 855-864
11. Unkles, S.E., Hawker, K.L., Grieve, C., Campbell, E.I., Montague, P., and Kinghorn, J.R. (1991) *crnA* encodes a nitrate transporter in *Aspergillus nidulans*. *PNAS* **88**, 204-208, (1995) erratum in *PNAS* **92**, 3076
12. Trueman, L.J., Richardson, A., and Forde, B.G. (1996) Molecular cloning of higher plant homologues of the high-affinity nitrate transporters of *Chlamydomonas reinhardtii* and *Aspergillus nidulans*. *Gene*, **175**, 223-231
13. Forde, B.G. (2000) Nitrate transporters in plants: structure, function and regulation. *Biochim. Biophys. Acta* **1465**, 219-235
14. Pao, S.S., Paulsen, I.T., and Saier, M.H., Jr. (1998) Major facilitator superfamily. *Microbiol. Mol. Biol. Rev.* **62**, 1-34
15. Abramson, J., Smirnova, I., Kasho, V., Verner, G., Kaback, H.R., and Iwata, S. (2003) Structure and mechanism of the lactose permease of *Escherichia coli*. *Science* **301**, 610-615
16. Huang, Y., Lemieux, M.J., Song, J., Auer, M., and Wang, D.N. (2003) Structure and mechanism of the glycerol-3-phosphate transporter from *Escherichia coli*. *Science* **301**, 616-620
17. Yin, Y., He, X., Szewczyk, P., Nguyen, T., Chang, G. (2006) Structure of the multidrug transporter EmrD from *Escherichia coli*. *Science* **312**, 741-744
18. Dang, S., Sun, L., Huang, Y., Lu, F., Liu, Y., Gong, H., Wang, J., Yan, N. (2010) Structure of a fucose transporter in an outward-open conformation *Nature* **467**, 734-738
19. Newstead, S., Drew, D., Cameron, A.D., Postis, V.L., Xia, X., Fowler, P.W., Ingram, J.C., Carpenter, E.P., Sansom, M.S., McPherson, M.J., Baldwin, S.A., and Iwata, S. (2011) Crystal structure of a prokaryotic homologue of the mammalian oligopeptide-proton symporters, PepT1 and PepT2. *EMBO J.* **30**, 417-426
20. Unkles, S.E., Rouch, D.A., Wang, Y., Siddiqi, M.Y., Glass, A.D., and Kinghorn, J.R. (2004) Two perfectly conserved arginine residues are required for substrate binding in a high-affinity nitrate transporter. *PNAS* **101**, 17549-17554

21. Jia, W., Tovell, N., Clegg, S., Trimmer, M., and Cole, J. (2009) A single channel for nitrate uptake, nitrite export and nitrite uptake by *Escherichia coli* NarU and a role for NirC in nitrite export and uptake. *Biochem. J.* **417**, 297-304
22. Clutterbuck, A.J. (1974) *Aspergillus nidulans*. In *Handbook of Genetics vol.1.* (King, R.C., ed) pp. 447-510, Plenum Press, New York
23. Cove, D.J. (1966) The induction and repression of nitrate reductase in the fungus *Aspergillus nidulans*. *Biochim. Biophys. Acta* **113**, 51-56
24. Unkles, S.E., Zhou, D., Siddiqi, M.Y., Kinghorn, J.R., and Glass, A.D. (2001) Apparent genetic redundancy facilitates ecological plasticity for nitrate transport. *EMBO J.* **20**, 6246-6255
25. Riach, M., and Kinghorn, J.R. (1995) in *Fungal Genetics* (Bos, C., ed) pp. 209-234, Wiley Press, London.
26. Katoh, K., and Toh, H. (2008) Recent developments in the MAFFT multiple sequence alignment program. *Brief. Bioinformatics* **9**, 286-298
27. Warrens, A.N., Jones, M.D., and Lechler, R.I. (1997) Splicing by overlap extension by PCR using asymmetric amplification: an improved technique for the generation of hybrid proteins of immunological interest. *Gene* **186**, 29-35
28. Southern, J.A., Young, D.F., Heaney, F., Baumgärtner, W.K., and Randall, R.E. (1991) Identification of an epitope on the P and V proteins of simian virus 5 that distinguishes between two isolates with different biological characteristics. *J. Gen. Virol.* **72**, 1551-1557
29. Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M. and Barton, G. J. (2009) Jalview Version 2 - a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189-1191
30. Brownlee, A.G., and Arst, H.N., Jr. (1983) Nitrate uptake in *Aspergillus nidulans* and involvement of the third gene of the nitrate assimilation gene cluster. *J. Bacteriol.* **155**, 1138-1146.
31. Kelley, L.A., and Sternberg, M.J.E. (2009) Protein structure prediction on the web: a case study using the Phyre server. *Nat. Protoc.* **4**, 363-371
32. Roy, A., Kucukural, A., and Zhang, Y. (2010) I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* **5**, 725-738
33. Smirnova, I., Kasho, V., Sugihara, J., Choe, J.Y., and Kaback, H.R. (2009) Residues in the H⁺ translocation site define the pKa for sugar binding to LacY. *Biochemistry* **48**, 8852-8860
34. Javadpour, M.M., Eilers, M., Groesbeek, M., and Smith, S.O. (1999) Helix packing in polytopic membrane proteins: role of glycine in transmembrane helix association. *Biophys. J.* **77**, 1609-1618
35. Holyoake, J., and Sansom, M.S. (2007) Conformational change in an MFS protein: MD simulations of LacY. *Structure* **15**, 873-884
36. D'Rozario, R.S, and Sansom, M.S. (2008) Helix dynamics in a membrane transport protein: comparative simulations of the glycerol-3-phosphate transporter and its constituent helices. *Mol. Membr. Biol.* **25**, 571-583
37. Lovell, S.C., Davis, I.W., Arendall, W.B., 3rd, de Bakker, P.I., Word, J.M., Prisant, M.G., Richardson, J.S., and Richardson, D.C. 2003 Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins* **50**, 437-450
38. Russ, W.P., and Engelman, D.M. (2000) The GxxxG motif: a framework for transmembrane helix-helix association. *J. Mol. Biol.* **296**, 911-919
39. Kleiger, G., Grothe, R., Mallick, P., and Eisenberg, D. (2002) GXXXG and AXXXA: common alpha-helical interaction motifs in proteins, particularly in extremophiles. *Biochemistry* **41**, 5990-5997
40. Wang, X., Ye, L., McKinney, C.C., Feng, M., and Maloney, P.C. (2008) Cysteine scanning mutagenesis of TM5 reveals conformational changes in OxlT, the oxalate transporter of *Oxalobacter formigenes*. *Biochemistry* **47**, 5709-5717
41. Enkavi, G., and Tajkhorshid, E. (2010) Simulation of spontaneous substrate binding revealing the binding pathway and mechanism and initial conformational response of GlpT. *Biochemistry* **49**, 1105-1114

42. Kim, S., Jeon, T.J., Oberai, A., Yang, D., Schmidt, J.J., and Bowie, J.U. (2005) Transmembrane glycine zippers: physiological and pathological roles in membrane proteins. *PNAS* **102**, 14278-14283
43. Rosenhouse-Dantsker, A., and Logothetis, D.E. (2006) New roles for a key glycine and its neighboring residue in potassium channel gating. *Biophys. J.* **91**, 2860-2873
44. Wang, X., Sarker, R.I., and Maloney, P.C. (2006) Analysis of substrate-binding elements in OxIT, the oxalate:formate antiporter of *Oxalobacter formigenes*. *Biochemistry* **45**, 10344-10350
45. Yang, Q., Wang, X., Ye, L., Mentrikoski, M., Mohammadi, E., Kim, Y.M., and Maloney, P.C. (2005) Experimental tests of a homology model for OxIT, the oxalate transporter of *Oxalobacter formigenes*. *PNAS* **102**, 8513-8518
46. Masterton, W.L., Bolocofsky, D., and Lee, T.P. (1971) Ionic radii from scaled particle theory of the salt effect *J. Phys. Chem.* **75**, 2809-2815
47. Lagerstedt, J.O., Voss, J.C., Wieslander, A., and Persson, B.L. (2004) Structural modeling of dual-affinity purified Pho84 phosphate transporter. *FEBS Lett.* **578**, 262-268
48. Sigal, N., Vardy, E., Molshanski-Mor, S., Eitan, A., Pilpel, Y., Schuldiner, S., and Bibi, E. (2005) 3D model of the *Escherichia coli* multidrug transporter MdfA reveals an essential membrane-embedded positive charge. *Biochemistry* **44**, 14870-14880
49. Perry, J.L., Dembla-Rajpal, N., Hall, L.A., and Pritchard, J.B. (2006) A three-dimensional model of human organic anion transporter 1: aromatic amino acids required for substrate transport. *J. Biol. Chem.* **281**, 38071-38079
50. Soares-Silva, I., Sá-Pessoa, J., Myrianthopoulos, V., Mikros, E., Casal, M., and Diallinas, G. (2011) A substrate translocation trajectory in a cytoplasm-facing topological model of the monocarboxylate/H⁺ symporter Jen1p. *Mol. Microbiol.* **81**, 805-817
51. Courville, P., Quick, M., and Reimer, R.J. (2010) Structure-function studies of the SLC17 transporter sialin identify crucial residues and substrate-induced conformational changes. *J. Biol. Chem.* **285**, 19316-19323

FOOTNOTES

This work was supported by a BBRSC award to SEU. VFS wishes to acknowledge support of a BBSRC studentship and NA the support of an HEC award from the Pakistan Government and the University of the Punjab.

FIGURE LEGENDS

FIGURE 1. Secondary structure model showing the position and sequence of the nitrate signature motifs within TM5 and TM11. Solid vertical bars represent the remaining ten TMs. Residues highlighted in grey in TM5 and TM11 form the core region of the signatures.

FIGURE 2. Amino acid sequences of the nitrate signatures and surrounding residues. Sequence logos for the 18 residue stretch encompassing NS1 and NS2 from prokaryotes (NS1pro and NS2pro, respectively) and eukaryotes (NS1euk and NS2euk, respectively), and the sequence of the corresponding regions in *A. nidulans* NrtA with numbers below referring to the NrtA amino acid sequence.

FIGURE 3. Models of NrtA in different conformations. Based on inward-facing GlpT (PDB ID: 1PW4) (*A*, *C* and *E*) and the outward-facing FucP (PDB ID: 3O7Q) (*B*, *D* and *F*). *A-D*, NrtA viewed from the plane of the membrane; *C*, a 180° rotation of *A*; *D* a 180° rotation of *B*. *E*, inward-facing GlpT-based model viewed from the outside looking towards the cytoplasmic side and *F*, outward-facing FucP-based model viewed from the cytoplasmic side looking toward the outside. Conserved glycine residues and S161 are represented by spheres, orange for TM5 and yellow for TM11. Side chains of residues R87, R368, N168 and N459 are shown as spheres in *A-D* and sticks in *E* and *F* with the carbon side chain

atoms coloured orange for R368 and N168, and yellow for R87 and N459. Residue G167 is hidden behind N168 in *A* and is not labelled. Residues labelled in red show very poor or no expression in western blots (see Fig. 4).

FIGURE 4. Western blot analysis of mutants. NrtA protein was detected in immunoblots of approximately 50 µg of protein from membrane fractions of the mutant strains indicated, grown at 37 °C. *A*, the NS1 region of TM5 and *B*, the NS2 region of TM11. Also included in the left of *A* and *B* is the wild type strain T454 and in *B*, the double *nrtA nrtB* deletion mutant T110. Protein was detected via the C-terminal V5 tag using anti-V5-horseradish peroxidase antibody.

Table 1. NS1 and NS2 mutant growth characteristics and net nitrate transport

Nitrate Signature 1			Nitrate Signature 2		
Mutant	Growth on nitrate	Net nitrate uptake (nmol/min/mg)	Mutant	Growth on nitrate	Net nitrate uptake (nmol/min/mg)
G157A	-	N.d.	G448A	++	4.4 ± 0.3
T158A	+++	8.4 ± 0.4	I449A	+++	8.9 ± 0.5
			V450A	+++	8.4 ± 0.5
N160A	++	4.8 ± 0.1	S451A	+++	9.1 ± 0.7
S161A	+++	9.5 ± 0.1	G452A	+	2.7 ± 0.1
L162A	+++	9.2 ± 0.4	M453A	+++	8.1 ± 0.7
			V454A	+++	7.7 ± 0.6
G165A	-	0.3 ± 1.0	G455A	+ / +++	3.5 ± 0.3
L166A	+ / +++	3.1 ± 0.4	G456A	+ / +++	2.6 ± 0.4
G167A	-	N.d.	F457A	+ / +++	2.7 ± 0.2
N168A	-	N.d.	G458A	+	4.7 ± 0.5
			N459A	-	N.d.
G170A	-	N.d.	L460A	+	4.2 ± 0.2
G171A	+	0.6 ± 0.4	G461A	- / +	1.2 ± 0.6
G172A	-	N.d.	G462A	+	4.1 ± 0.3
I173A	+++	7.9 ± 0.6	I463A	+++	7.9 ± 0.1
T174A	+++	7.4 ± 0.2	I464A	+++	9.0 ± 0.9
			F465A	+++	8.4 ± 0.8
T454					
	+++	8.7 ± 0.3			
T110					
	-	0.1 ± 0.9			

Growth tests for nitrate were performed on minimal medium (pH 6.5) with 1 mM, 10 mM or 100 mM nitrate as the sole source of nitrogen for 2 days at 37 °C. The relative pattern of growth was the same at all concentrations of nitrate. Growth responses were scored as - no growth, + slight growth, ++ intermediate growth, +++ wild-type growth. A similar pattern of growth was exhibited at 25 °C except for the mutants G171A in NS1 and G448A and G455A in NS2, where growth approximating to wild type level was restored. Net nitrate uptake experiments were performed only on strains that produced appreciable growth on 100 mM nitrate (+ to +++). N.d. denotes assays not performed on strains that failed to show any growth on nitrate. Young conidiospores were grown for uptake assays to germination in 5 mM urea as the sole nitrogen source and induced with 10 mM nitrate for 100 min. All strains harboured a deletion mutation in the other *A. nidulans* gene (*nrtB*) for nitrate transport (24). Control strains T454, wild type *nrtA* integrated at *argB*, and T110, double deletion mutant for *nrtA* and *nrtB* are included.

Figure 1

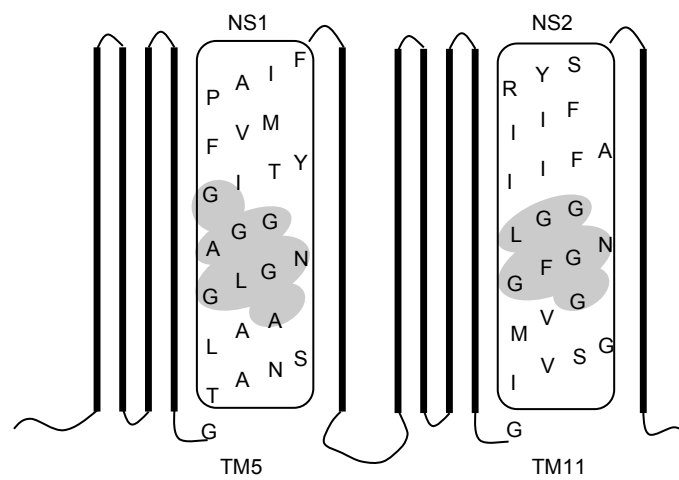


Figure 2

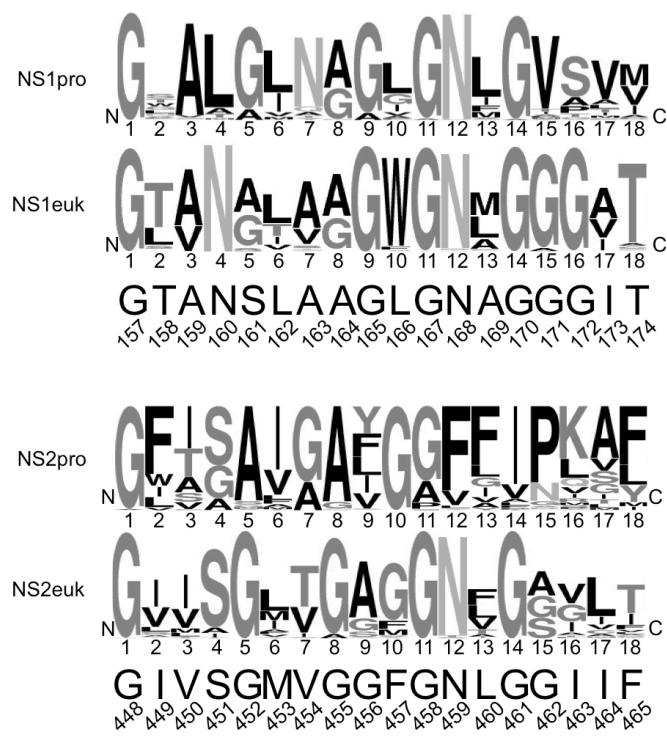


Figure 3

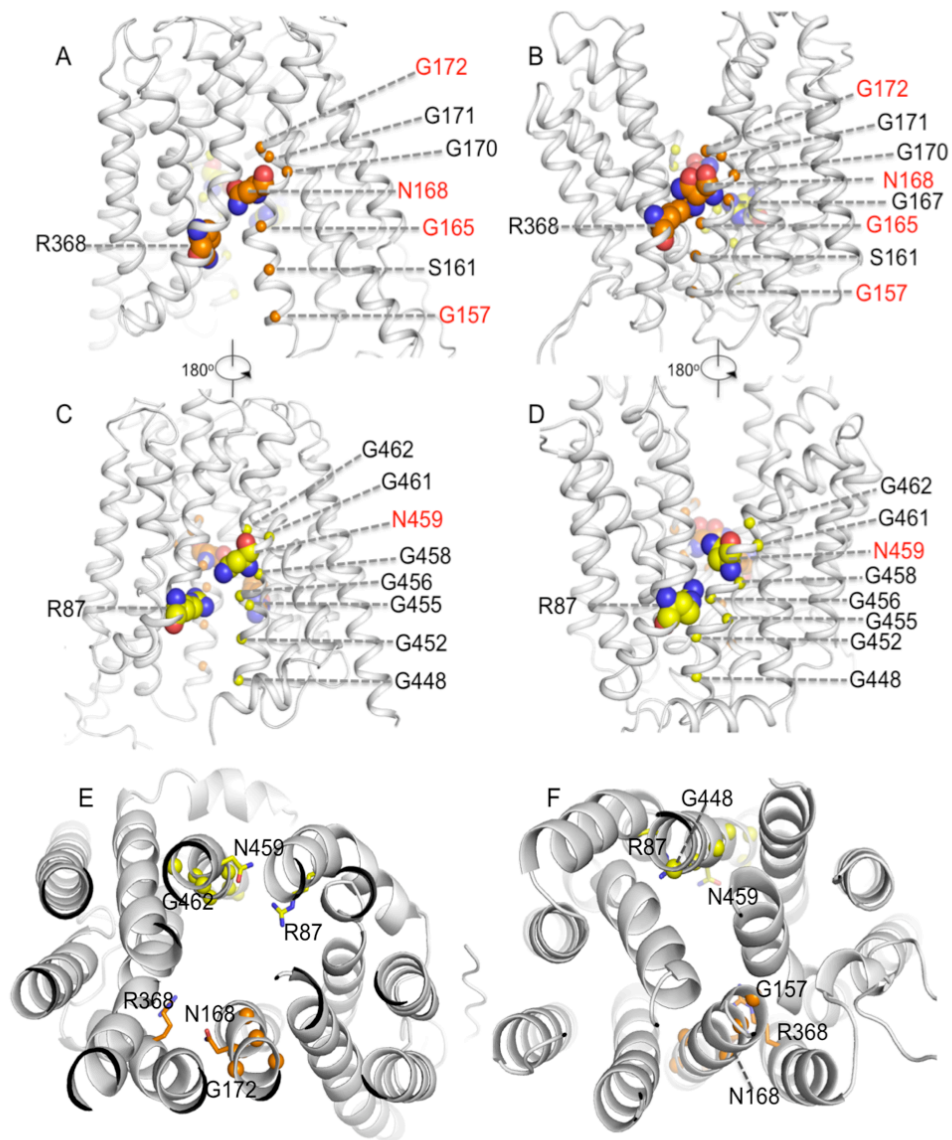


Figure 4

